基于区块链和隐私计算的联邦学习数据安全聚合方法

孙弘业，王冬宇

北京邮电大学人工智能学院，北京 100876

摘要： 随着 5G 与人工智能的发展，海量分布式智能终端生成的本地数据呈指数级增长。为在保护用户隐私的同时促进数据共享，联邦学习技术通过参数聚合实现了分布式模型训练。然而，传统联邦学习的中心化架构存在单点故障风险，且面临推理攻击等隐私泄露威胁。为此，本文提出一种融合区块链与隐私计算的数据安全聚合策略，解决跨区域多终端场景下模型参数传输与聚合中的隐私泄露与数据完整性问题。该方案搭建了以区域服务器簇为核心的联邦学习架构，通过分布式存储与计算降低单点故障风险，并结合区块链智能合约确保数据聚合的透明性、可追溯性与不可篡改性。进一步提出轻量化隐私保护方案，结合椭圆曲线同态加密与动态门限 Shamir 秘密共享技术，实现梯度参数的分片加密上传与分布式聚合。理论分析表明，该方案在半诚实与主动攻击模型下可有效抵御梯度反演与合谋攻击。实验证明，其在模型精度损失可控的前提下，显著降低了通信与计算开销，验证了其实际应用的可行性与高效性。

关键词： 联邦学习，区块链，隐私计算，数据聚合，Shamir 秘密共享，同态加密

A Federated Learning Data Security Aggregation Approach based on Blockchain and Privacy Computing

SUN Hong-Ye, WANG Dong-Yu

School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing 100876

Abstract: With the development of 5G and artificial intelligence, the local data generated by massive distributed intelligent terminals is growing exponentially. In order to promote data sharing while protecting user privacy, federated learning technology realizes distributed model training through parameter aggregation. However, the centralized architecture of traditional federated learning has the risk of single point of failure, and faces privacy disclosure threats such as inference attacks. Therefore, this thesis proposes a data security aggregation strategy that integrates blockchain and privacy computing to solve the privacy leakage and data integrity problems in the transmission and aggregation of model parameters in cross-region multi-terminal scenarios. The solution builds a federated learning architecture

作者简介：孙弘业（1999-），男，硕士研究生，主要研究方向：联邦学习。通信作者：王冬宇（1984-），男，副教授，通信方式：dy_wang@bupt.edu.cn，主要研究方向：无线通信，区块链。

with regional server clusters as the core, reduces the risk of single point of failure through distributed storage and computing, and ensures transparency, traceability and immutability of data aggregation combined with blockchain smart contracts. Further, a lightweight privacy protection scheme is proposed, which combines elliptic curve homologous encryption and dynamic threshold Shamir secret sharing technology to realize fragment encryption upload and distributed aggregation of gradient parameters. Theoretical analysis shows that this scheme can effectively resist gradient inversion and collusion attacks under semi-honest and active attack models. The experimental results show that the communication and calculation cost are significantly reduced under the premise of controllable model accuracy loss, and the feasibility and high efficiency of the practical application are verified.

Key words: federated learning, blockchain, privacy computing, data aggregation, Shamir secret sharing, homomorphic encryption

0 引言

联邦学习 (Federated Learning, FL) 技术 [1] 的提出有效地解决数据隐私保护与数据共享价值之间矛盾。作为一种革新性的分布式机器学习方法，联邦学习能够在多个参与者之间有效地共享知识而无需直接交换原始数据，可用于在持有不同本地数据集的用户之间训练人工智能模型，从而在保护个人数据隐私的同时实现数据共享的价值最大化。然而，联邦学习在隐私保护方面仍面临挑战。首先，传统 FL 的集中式服务器架构的安全仍存在一定隐患，一旦中央服务器遭受攻击，可能导致系统瘫痪、用户隐私泄露或模型训练被篡改；其次，尽管联邦学习采用传输模型参数，而非直接传递明文数据的方法来保障本地隐私的安全，然而它仍然面临诸如隐私推理攻击等安全威胁。

为应对上述问题，本文基于区块链和隐私计算技术提出一种联邦学习数据安全聚合策略，通过搭建去中心化的区块链网络架构，并结合秘密共享和同态加密方法，确保模型参数在传输和聚合过程中的隐私保护，旨在形成具有鲁棒性和更高效的隐私保护机制，以提升联邦学习系统的数据隐私安全。

1 系统模型总体设计

本节提出了一个针对跨区域多终端联邦学习场景的网络模型，称为隐私聚合链（Privacy Aggregation Chain, PA-Chain）。该模型在分布式双层区块链网络架构的基础上引入了区域服务器簇，并通过区域主服务器作为区块链节点参与共识和全局聚合。PA-Chain 利用区块链技术实现了加密模型的传输和安全聚合，进一步增强了系统的分布式信任机制和抗攻击能力。同时，为了结合秘密共享和同态加密技术，本节优化设计了区块结构。为了应对可能的安全威胁，本节构建了一个全面的威胁模型，包括半诚实服务器模型、主动恶意窃取以及合谋攻击等攻击模型手段，为后续数据安全聚合策略的设计明确了应对目标。

1.1 系统架构

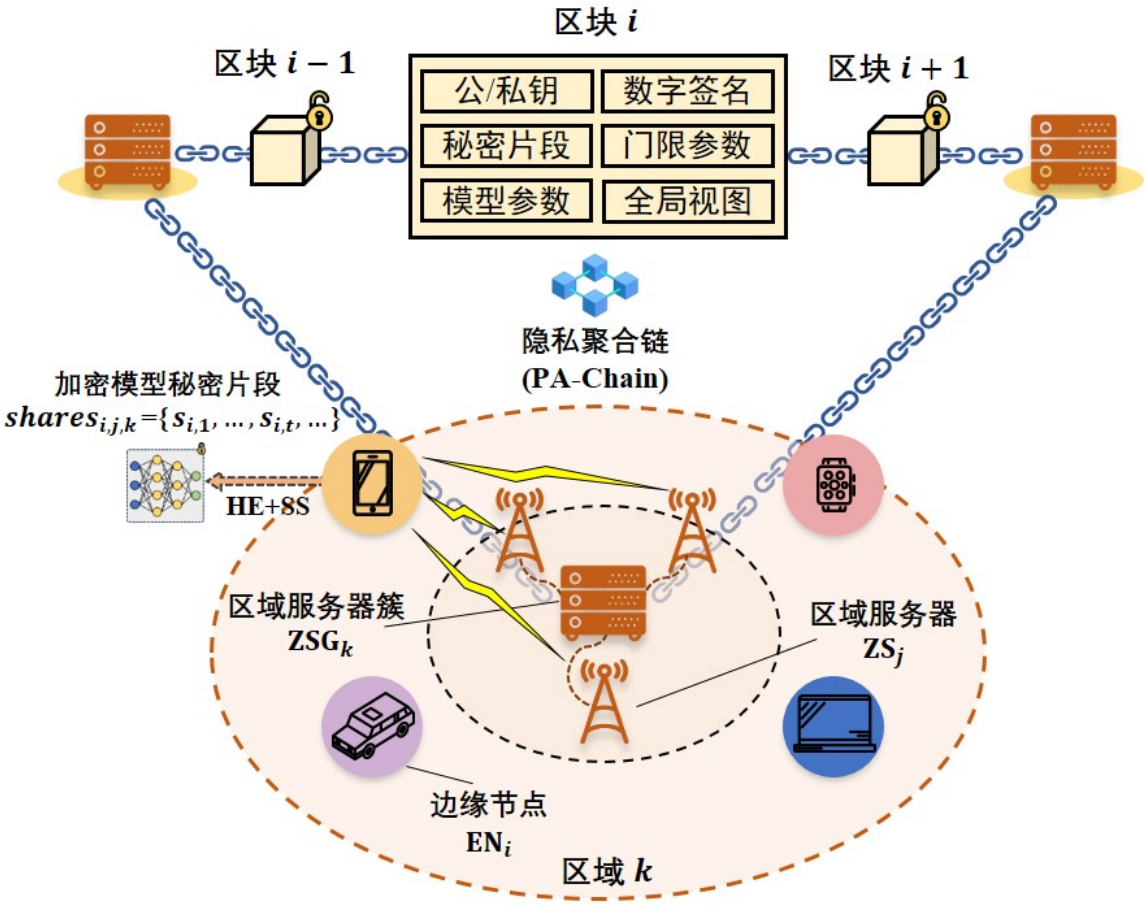

图 1: 基于区块链的联邦学习数据安全聚合网络架构（PA-Chain）

在基于课题 1 定义的跨区域多终端网络环境的基础上，本节进一步搭建了基于区块链的联邦学习数据安全聚合网络架构，其中区块链网络称为隐私聚合链，如图 1所示。该架构通过引入区域服务器簇（Zone Server Group, ZSG）的角色支持秘密共享技术的实现。具体地，PA-Chain 网络架构由区块链网络层、区域服务器簇层和边缘节点层三层组成，分别承担不同的功能与责任：

- 边缘节点层：边缘节点（Edge Node, EN）即为联邦学习中的参与者，通常是移动设备或其他用户终端。每个边缘节点在本地数据集上进行模型训练，生成本地模型参数，并将这些模型参数经过同态加密和 Shamir 秘密共享后形成秘密片段，并分别发送至附近的区域服务器簇中的各个 ZS 处。由于每个边缘节点的模型参数在上传前会被加密和分片，即便攻击者捕获了某个片段，也无法恢复完整的模型参数，从而有效保障了数据的隐私性。
- 区域服务器簇：ZSG 在边缘节点层之上，由一个区域主服务器协同多个具备一定计算能

力和通信资源的区域服务器（Zone Server, ZS）组成。区域服务器簇在上传过程中充当中介角色，负责接收、处理并加密边缘节点上传的模型参数。具体地，对于服务区 k 的 $\mathrm{ZSG}_k = \{\mathrm{ZS}_{0,k}; \mathrm{ZS}_{1,k}, \ldots, \mathrm{ZS}_{m,k}\}$, $\mathrm{ZS}_{0,k}$ 表示区域主服务器，$\mathrm{ZS}_{m,k}$ 表示与区域主服务器有线连接的 m 个区域服务器。ZSG 中的每个 ZS 接收到的模型片段在分片前均经过加密处理，且通过 Shamir 秘密共享方式将加密模型分割成多个秘密片段，存储在不同的区域服务器中，以避免单点故障或攻击。区域主服务器负责收集各 ZS 接收到的秘密片段并进行模型重构，然后完整的区域加密模型参数上传至区块链。在模型下载阶段，ZSG 从区块链中获取聚合后的全局模型，并将其分发给所在区域的边缘节点，供用户进行下一轮次的本地训练。

- 区块链网络层：区块链网络是 PA-Chain 架构中的核心部分，负责管理整个系统的隐私计算机制和数据聚合过程。区块链用于记录和验证模型参数的上传与聚合，在该层，用户和边缘节点的注册、身份验证以及模型参数的聚合都通过智能合约进行管理。区块链的不可篡改性保证了数据的完整性，而智能合约机制确保了模型聚合过程的透明性和自动化。同时，区块链网络通过共识控制各 ZSG 动态更新每轮的同态加密公私钥和秘密共享门限来灵活调整每轮的加密聚合策略，以实现通信资源消耗和数据隐私保护的平衡。

整个 PA-Chain 架构通过上述三层分布式网络的协同工作，保障了模型数据的隐私性、安全性和高效性。

1.2 区块链设计

PA-Chain 的区块结构设计充分考虑了架构中各层之间的数据传递、隐私保护以及动态加密参数的更新需求。该区块结构主要分为区块头和区块体两大部分，每一部分又包含若干字段，具体结构设计如图 2所示。

（1）区块头：每个区块的元数据，起到链接前后区块、记录全局状态及动态参数更新信息的作用，主要包含以下字段：

- 版本号（Version）：表示区块协议的版本信息，确保所有参与节点对区块结构和验证规则达成一致，有助于后续协议升级和兼容性维护。
- 前一区块哈希（Previous Block Hash）：存储前一区块头的哈希值，实现区块链的串联结构，保证区块数据的不可篡改性和链式完整性。
- 时间戳（Timestamp）：记录区块生成的精确时间，既用于防止重放攻击，也有助于后续模型聚合轮次的时间调度与溯源。
- 轮次标识（Round Identifier）：指明当前区块对应的联邦学习训练轮次，便于各参与方同步模型更新进程。
- Merkle 根（Merkle Root）：通过对区块体中所有交易（例如各区域服务器聚合后的模型上传记录）进行哈希构建得到的 Merkle 树根，确保区块体数据完整性及快速验证。

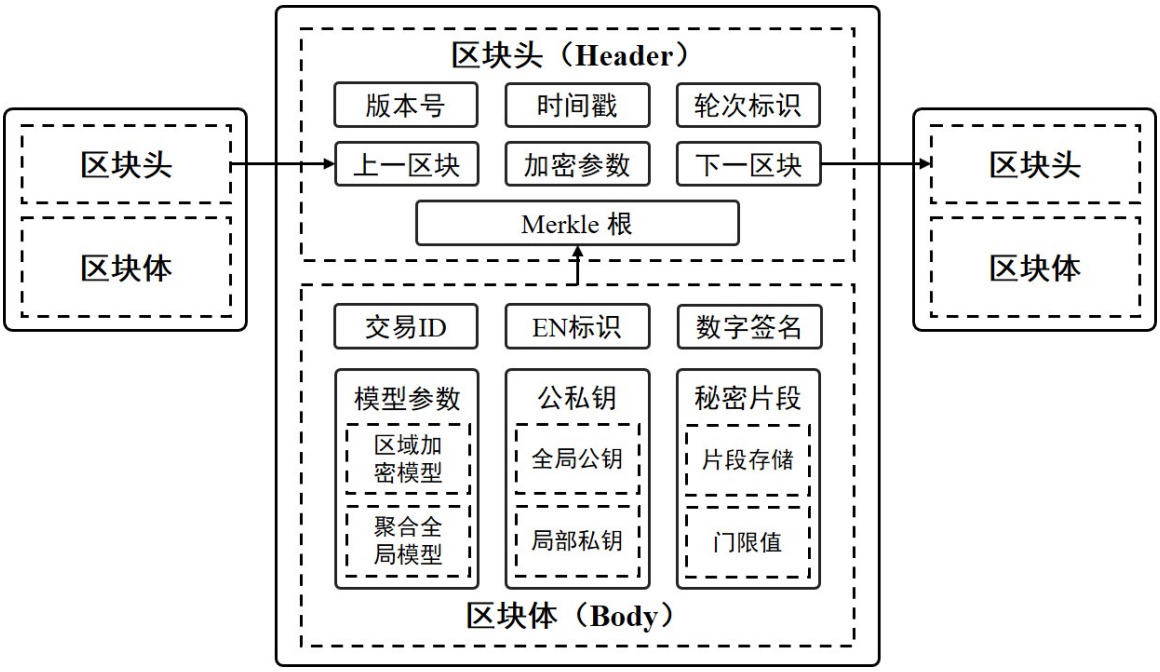

图 2: PA-Chain 的区块结构示意图

- 加密参数信息（Encryption Parameters）：记录本轮联邦学习所采用的同态加密公私钥信息以及 Shamir 秘密共享门限参数。这一字段由区块链通过共识机制动态更新，确保每一轮聚合均使用最新的加密策略，从而在保证数据隐私的同时平衡通信资源消耗。

（2）区块体：主要用于存储经过加密处理并经过区域服务器簇聚合后的模型数据上传记录。每个区块体内包含多个交易记录，每笔交易对应一个区域主服务器提交的聚合结果。除了包含交易 ID、发送方标识和时间戳等常规信息存储区域外，还应包含以下设计部分：

- 加密模型参数：该字段存放经过同态加密处理的区域级模型参数。由于原始模型参数经过区域服务器簇中各个 ZS 间采用 Shamir 秘密共享分片后重构而来，因此即使部分数据泄露，也无法恢复出完整模型，保障数据隐私性。

- 数字签名及验证信息：用于证明交易数据的真实性与完整性。每个交易附带数字签名，确保数据在传输过程中未被篡改，并可通过区块链中预先注册的身份信息进行验证。

- 同态加密公私钥：PA-Chain 网络的核心组成部分之一就是管理同态加密公私钥的区域。在每一轮训练过程中，区块链会通过智能合约控制公私钥的生成、分发和更新，确保参与方使用最新的加密参数进行加密和解密操作。这些密钥信息会被保存在区块链上，即分布在 ZSG 的区域中央服务器和分布式 ZS 中，以便在数据上传和模型聚合时进行验证和解密。

- 秘密片段存储区：ZSG 中的每个区域服务器在接收和处理边缘节点上传的秘密片段时，每个区块中包含一项存储分片的记录，详细记录该分片存储在哪些具体的 ZS 上。通过智能

合约，PA-Chain 管理区域服务器与模型分片之间的映射关系。每个区域服务器的加密模型片段信息将被存储在区块链上，并通过智能合约定期更新和发送至区域主服务器上。当模型需要聚合时，智能合约通过这些映射关系帮助区域主服务器重构节点的完整模型。

通过上述区块结构设计，PA-Chain 不仅实现了模型数据的安全、不可篡改存储，还通过区块头中动态更新的加密参数（同态加密公私钥与秘密共享门限信息）支持每轮联邦学习中的加密聚合策略灵活调整。同时，区块体中对每轮上传数据详细记录了区域服务器身份信息、加密后的模型参数及边缘节点数字签名信息，确保数据隐私和完整性得到多重保障。该设计为后续秘密共享与同态加密技术的具体实现提供了技术基础。

1.3 威胁模型与设计目标

基于区块链的联邦学习数据安全聚合网络架构的提出旨在保护边缘节点在上传本地模型参数的过程中隐私泄露的风险。因此，各边缘节点和区域服务器簇之间存在恶意行为或相互串通合谋的可能性，具体威胁模型设置的规定如下：

（1）半诚实服务器模型：假设区域服务器簇中包括主服务器在内的所有区域服务器都是诚实的，但可能会对其他节点进行某些恶意行为或窥探其信息。这些服务器会按照协议执行操作，但它们可能会根据协议执行之外的其他方式来提取或分析敏感数据。具体地，半诚实服务器不会篡改或伪造数据，完全按照既定的规则与边缘节点通信并收集和聚合局部模型，但它们会在模型训练过程中试图从加密的数据或已知的数据片段中推断出其他信息。例如，区域服务器可能通过观察或分析边缘节点上传的模型更新来推测原始数据内容，从而泄露数据隐私。

（2）主动恶意攻击模型：与半诚实模型相比，主动恶意节点模型考虑了完全不可信的区域服务器或边缘节点，这种攻击模型包括推理攻击和中间人攻击：

- 推理攻击：恶意区域服务器或恶意窃听节点可能通过推理攻击 [2] 来恢复或推测其他边缘节点的私有数据。例如，攻击者可能从边缘节点上传的模型参数中推测出用户的敏感数据，或基于模型的某些特性（如梯度、权重等）推测出其他节点的训练数据。

- 中间人攻击：攻击者可能通过拦截或修改在区域服务器和边缘节点之间传输的数据包来获取敏感信息，或操控数据传输流程 [3]。例如，攻击者可能伪装成通信中的一方，欺骗对方信任其所发送的数据。攻击者不仅能够看到通信内容，还能篡改数据或向通信中插入虚假信息，这可能导致重要数据的泄露或操作的失败。

基于以上网络架构和威胁模型，为了保证所提出方法的性能，这里定义了方案的四个设计要求：

（1）安全性：节点的本地数据集和从中计算的本地梯度应该对所有服务器和其他参与者保密。任何其他参与方都无法从联邦学习系统中传递的任何消息中获取私人信息，因为这些消息会泄露参与者的本地数据和模型。

（2）非协作性：在该方案的训练过程中，边缘节点无需相互之间进行直接通信，只需与服务器进行数据交换，从而降低了模型训练过程中的通信复杂度。

（3）准确性：隐私保护方案应尽量减少对全局模型准确性的负面影响。

（4）高效性：隐私保护方案所带来的额外计算和通信成本应尽可能小，尤其是参与者所承担的成本。

2 基于隐私计算的数据动态安全聚合策略

2.1 基于 ECC-Paillier 的轻量化同态加密聚合方法

同态加密技术是隐私保护联邦学习的核心支撑，其中 Paillier 加密方案 [4] 因其加法同态特性成为经典方法。传统 Paillier 算法在有限域上构建加密系统，其核心优势在于满足：

$$E(m_1) \cdot E(m_2) = E(m_1 + m_2) \mod n^2 \tag{1}$$

其中 $n = pq$ 为大素数乘积，该性质使得云端可直接对密文执行聚合运算而无需解密。然而，Paillier 方案存在两个显著缺陷：其一，密钥尺寸过大（2048-bit 公钥导致密文膨胀至 4KB/参数）；其二，大整数模幂运算复杂度高达 $O(log^3 n)$。在跨区域多终端场景下，边缘设备受限于计算能力与网络带宽，上述缺陷将导致以下瓶颈：（1）通信效率低下：移动网络传输大量密文梯度（如 ResNet-152 含千万级参数）时，单轮通信延迟超过分钟级；（2）终端能耗过高：低功耗设备执行 Paillier 加密的 CPU 峰值功耗可达 1.2W，难以支持持续训练；（3）扩展性受限：大规模节点参与时，密钥管理与分发面临显著协调开销。为此，本节提出基于椭圆曲线密码（Elliptic Curve Cryptography，ECC）优化的 ECC-Paillier 混合加密方案 [5]。该方法通过将 Paillier 映射到椭圆曲线离散对数问题（Elliptic Curve Discrete Logarithm Question, ECDLP）域，结合椭圆曲线的离散对数问题（ECDLP）与 Paillier 的判定性复合剩余假设（DCR），构建安全高效的加密框架并实现三重优化：

（1）密文压缩：将原方案中 $\mathbb{Z}_{n^2}^*$ 上的运算迁移至椭圆曲线群 $E(\mathbb{F}_p)$，使得同等安全强度下密钥长度缩减至 256-bit；

（2）计算加速：椭圆曲线标量乘法复杂度降为 $O(\log k)$，较 Paillier 模指数运算提升 4-5 倍效率。ECC-Paillier 和 Paillier 方法的效率对比曲线如图 3 所示。

（3）动态适配：支持在素数域 $y^2 = x^3 + ax + b \mod p$ 与二进制域 $\mathrm{GF}(2^m)$ 间灵活切换，适配异构终端硬件。

具体地，设椭圆曲线参数为 (a, b, p, G, n, h)，其中 G 为基点，n 为曲线阶，h 为余因子，算法原理如下：

- 密钥生成: 选择随机私钥 $sk \in [1, n-1]$，计算公钥 $pk = sk \cdot G \in E$;
- 加密过程: 对明文矩阵 $\Delta g \in \mathbb{Z}_p$, 生成随机数 $\pi \leftarrow \mathbb{Z}_n$, 计算:

$$E(\Delta g) = (\pi \cdot G, \Delta g \cdot H + \pi \cdot pk) \tag{2}$$

其中 $H = \mathrm{Hash}(G) \in E$ 为哈希映射点。

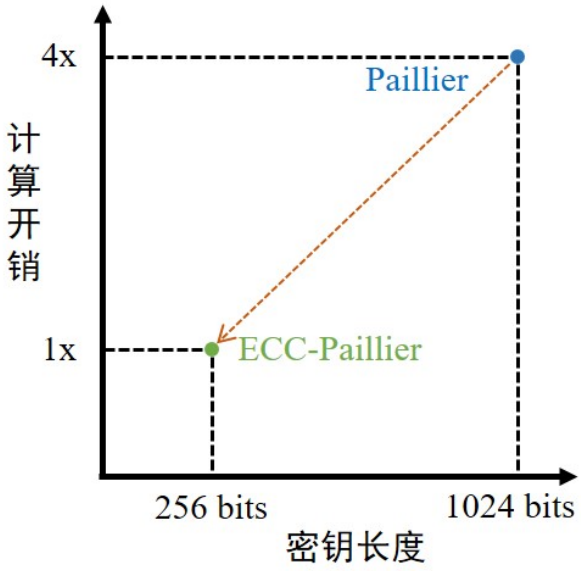

图 3: ECC-Paillier 和 Paillier 方法的效率对比曲线

- 同态加法: 对两密文 $C_1 = (P_1, Q_1)$ 和 $C_2 = (P_2, Q_2)$，聚合结果为:

$$C_{\text{sum}} = (P_1 + P_2, Q_1 + Q_2) \tag{3}$$

- 解密过程: 使用私钥 sk 恢复明文:

$$\Delta g = (Q - sk \cdot P) \cdot H^{-1} \tag{4}$$

针对一个包括 l 个 ZSG 的 PA-Chain 网络，其中对于包含 m 个 ZS 的序号为 k 的区域中的 EN_i，规定其基于本地数据生成的局部模型参数 $g_{i,k}$。EN_i 需要确保其在上传和聚合期间保持加密状态。PA-Chain 场景下 ECC-Paillier 同态加密数据聚合方案的流程如图 4所示，具体步骤如下所述：

（1）分布式密钥生成

由所有 ZSG 的区块链节点（区域主服务器）组成的区块链网络通过共识协议选出当前轮次的领导者负责全局模型聚合，并执行分布式密钥生成（distributed key generation protocol, DKG）协议：

$$(pk, sk_1, \ldots, sk_k, \ldots sk_l) \leftarrow \text{DKG}(l, t) \tag{5}$$

其中 sk_k 为私钥片段，需 $t+1$ 个节点联合才能解密（阈值加密）。

（2）密钥广播

公钥 pk 通过区块链智能合约广播至所有 ZSG，并通过分散的区块服务器下发至参与训练的 EN 节点，包含记录公钥哈希值确保完整性。

（3）局部模型加密：HE.Enc()

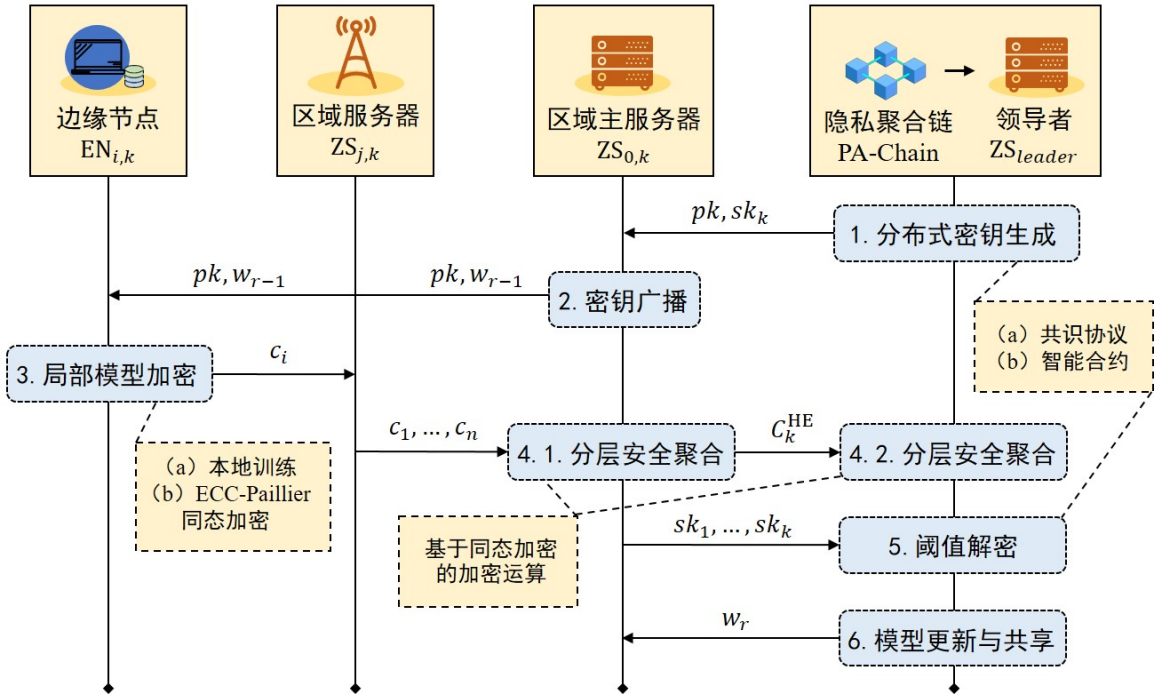

图 4: 基于 ECC-Paillier 的轻量化同态加密聚合方法流程

EN_i 使用文献 [5] 中研究的 ECC-Paillier 算法加密本地模型如公式 6 所示。

$$c_i = E(pk, \Delta g_i) = (\pi_i G, \Delta g_i \cdot H + \pi_i \cdot pk) \tag{6}$$

其中 $\pi_i \in \mathbb{Z}_p^*$ 为随机数，$H = \mathrm{Hash}(G) \in E$。

（4）分层安全聚合：HE.add()

ZS 接收 EN 的密文模型并进行验证，上载至主服务器，由主服务器对合法模型进行第一阶段区域聚合，即 ZS_j 执行密文加法聚合:

$$C_k^{\mathrm{HE}} = \bigoplus_{i \in S_k} c_i = \left(\sum_{i \in S_k} \pi_i G, \sum_{i \in S_k} (\Delta g_i \cdot H + \pi_i \cdot pk) \right) \tag{7}$$

其中 C_k^{HE} 表示区域主服务器 $ZS_{0,k}$ 计算的区域聚合结果，S_k 表示区域 k 内的 EN 节点集合，$\bigoplus$ 为同态加法聚合计算。

当所有区域服务器第一阶段聚合后，触发第二阶段全局聚合的智能合约，Leader 收集所有共识节点的区域模型，并加密计算全局聚合结果：

$$C_{\mathrm{HE}}^{global} = \bigoplus_{k=1}^{l} \mu_k \cdot C_{\mathrm{HE}}^k, \quad \mu_k = \frac{|S_k|}{N} \tag{8}$$

（5）解密请求：HE.Dec()

Leader 发起解密请求，至少 $t+1$ 个共识节点协作解。这里使用拉格朗日插值法进行解密，分为密钥征集（HE.Dec.select()）和阈值解密（HE.Dec.threshold()）两个阶段：

$$s = \prod_{k \in T} sk_k^{\lambda_k} \mod q \tag{9}$$

$$\Delta G = (C_2^{global}) \cdot (C_1^{global})^{-s} \mod p^2/N \tag{10}$$

其中 λ_k 为插值系数，q 为椭圆曲线阶数，s 是使用拉格朗日插值进行计算的系数，ΔG 为更新的全局梯度参数。

（6）模型更新验证和共享

通过区块链智能合约验证模型更新有效性，随后将全局模型明文 w_r 下发到所有 ZSG 中，完成一轮的 FL 训练。

由此，上述方法实现了在保护边缘节点和服务器之间传输的局部模型树度隐私的同时，支持密文状态下的安全聚合，确保全局模型的更新过程无泄露敏感数据。此外，ECC-Paillier 通过椭圆曲线优化了计算开销和通信成本，使其更适用于资源受限的跨区域终端场景，同时增强了系统的扩展性和抗攻击能力。

2.2 基于门限 Shamir 秘密共享的模型加密聚合机制

在 PA-Chain 规定的跨区域多终端联邦学习场景中，保护边缘节点与区域服务器层的安全通信及模型隐私保护是关键挑战。本节针对区域服务器簇提出一种基于动态门限 Shamir 秘密共享（Threshold Shamir Secret Sharing, TSSS）的模型加密聚合机制，解决本地模型分片传输、动态验证与安全聚合问题，重点保护隐私数据在边缘网络传输过程中的安全性。

2.2.1 Shamir 秘密共享基础原理

本论文采用了 Shamir 的 (n,t) -秘密共享方案 [6]。该方案允许用户将一个秘密 s 拆分成 n 个份额，使得任何 t 个份额可以重构出 s，而任何少于 t 个份额的组合无法获得关于 s 的任何信息。该方案由共享算法和重构算法组成。共享算法 $SS.share(s,t,n) \rightarrow \{s_1, s_2, \ldots, s_n\}$ 以一个秘密 s，限制和参数 t 为输入，生成一组份额 $\{s_1, s_2, \ldots, s_n\}$，该秘密 $s \in \mathbb{Z}_p$ 在有限域构造 t-次多项式时，满足以下公式：

$$f(x) = s + a_1 x + a_2 x^2 + \cdots + a_{t-1} x^{t-1} \tag{11}$$

其中 a_i 由伪随机函数生成，确保系数不可预测性。一个安全的伪随机生成器（Pseudorandom Generator, PRG）是一个确定性但不可预测的函数，它将一个二进制比特串扩展为一个更长的比特串。这里可以将 PRG 函数简化为 $\mathrm{PRG}(\{0,1\}^\lambda) \rightarrow \{0,1\}^{p(\lambda)}$，其中 $p(\lambda) > \lambda$。PRG 的输入是一个种子，即长度为安全参数 λ 的比特串。伪随机生成器的安全性保证只要种子保密，几乎不可能在计算上区分 PRG 的输出与真随机序列。当相同的种子输入到同一个伪随机生成器时输出是相同的，这减少了通信过程中的开销。

对于秘密的重构，选择一个互斥集 $x_1, x_2, \ldots, x_m$ 来计算分片 $s_i = f(x_i)$，则应该满足：

（1）门限特性：任何 t 个片段可以重构 s，少于 t 个片段不会泄露信息。

（2）同态性：分片间保持加法同态，支持密文域聚合。

即给定一个子集 R，其中 $R \subseteq S$ 且 $|R| > t$，重构算法 $SS.recon(R,t) \to s$ 可重构秘密 s。

2.2.2 动态 TSSS 联邦学习数据聚合协议

本节基于 TSSS 设计了动态联邦学习数据聚合协议，包括初始化、秘密片段生成与分发、分布式验证、安全重构与聚合共四个阶段，其协议框架和工作流程如图 5所示。

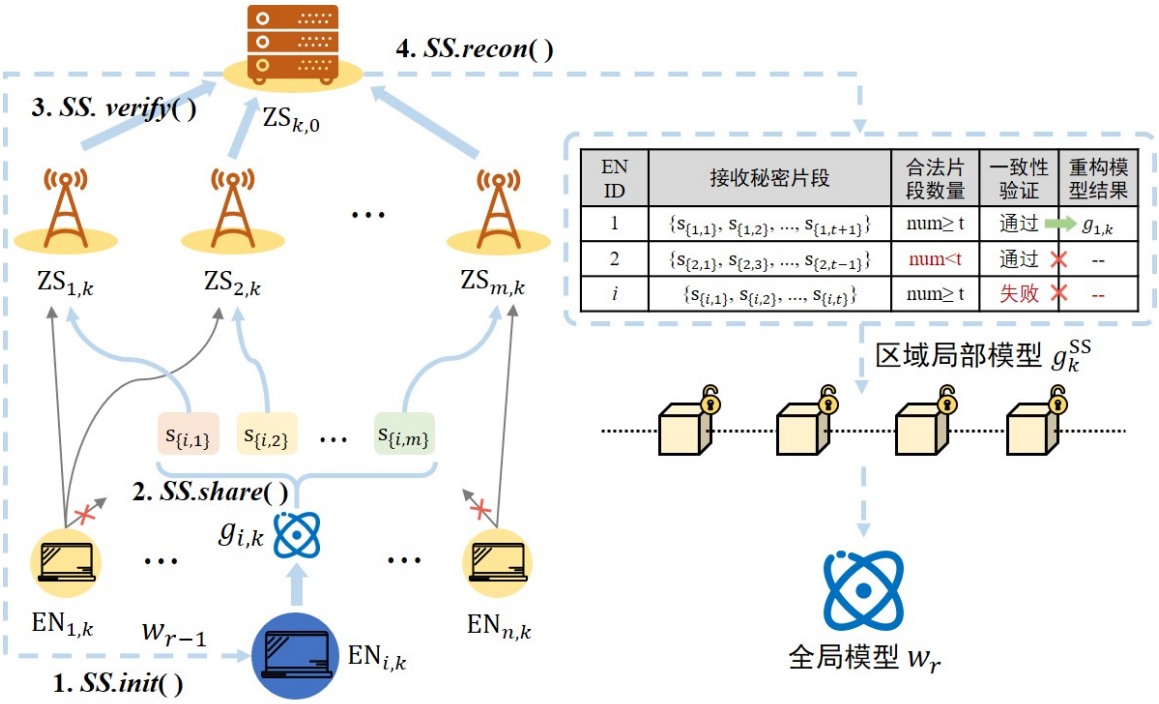

图 5: 基于动态门限秘密共享的数据聚合协议流程

（1）初始化阶段：SS.init()

每轮训练的初始，系统需配置和统一公共参数与动态门限参数，确保包括边缘节点、区域服务器和主服务器在内的所有参与方具备一致的初始环境。具体而言，ZSG 的主服务器在获取区域内所有 EN 的身份信息和 ZS 的活动状态后，需完成以下配置：选定一个足够大的素数域 GF(p)，其中 p 的位数需满足本地模型参数编码需求（例如 $p > 2^{256}$）；PRG 的种子 S_P 由主服务器秘密生成，用于后续多项式系数的动态构造；门限值 t 由主服务器根据当前轮次注册的 EN 数量 n 动态计算或处于工作状态的 ZS 数量 m，具体公式为：

$$t = \min(\lfloor \lambda n \rfloor + 1, m) \tag{12}$$

其中 λ 为预设的安全系数。此处 n 不会无限增大，通常会通过 FL 节点选择算法限制在一定范围内，故正常情况下 $\lfloor \lambda n \rfloor + 1 < m$；完成上述步骤后，主服务器向 ZSG 内所有 ZS 广播参数集

$\{\mathrm{GF}(p), S_P, \mathrm{Hash}(\cdot)$, 动态门限 t, ZS 活跃数 $m\}$，同时每个 EN 在注册后经由 ZS 同步获取这些参数以保障分片生成标准的一致性。

（2）秘密片段生成与分发：SS.share()

该阶段是秘密共享技术的核心步骤之一，EN 将本地模型参数安全分割并分发至通信覆盖能力内的所有 ZS 节点，以防止单点泄露。当 EN_i 生成本地梯度 $g_{i,k}$ 时，首先将其按维度映射至 $\mathrm{GF}(p)$ 域，随后利用公式 11 构造一个 $t-1$ 次多项式 $f_i(x) = g_i + a_1 x + a_2 x^2 + \cdots + a_{t-1} x^{t-1}$。其中，多项式系数 a_1 至 a_{t-1} 通过 $\mathrm{PRG}(S_P||\kappa)$ 生成，这里 κ 为系数索引值，旨在确保不同 EN 的系数独立且不可预测。针对每个 $\mathrm{ZS}_{j,k}$ $(j = 1, \ldots, m)$，EN 计算分片 $s(i,j) = f_i(j)$ 并附加时间戳 T_{ts} 与数字签名 $\sigma_i = \mathrm{Sign}(s(i,j)||T_{ts})$，防止重放攻击。分片通过点对点通信发送至对应的 ZS 节点，若检测到部分 ZS 失效，EN 可动态选择其他可用 ZS 补充分片以维持冗余。分片分发过程中，EN 需向主服务器预提交梯度 $g_{i,k}$ 的哈希承诺 $C^i_{hash} = H(g_i||r_i)$，其中 r_i 为随机抽值，用于后续一致性验证。

（3）分布式验证: SS.verify()

ZS 在接收秘密片段后执行两级验证机制。第一级验证由各 ZS 独立完成，包括检查分片 $s(i,j)$ 是否属于 $\mathrm{GF}(p)$ 域的有效元素，以及验证签名 σ_i 的合法性与时间戳的时效性。通过初步验证的秘密片段将被转发至主服务器，触发第二级验证流程。主服务器随机选择一组非零验证点 x^*（如 $x^* = m+1, m+2$ 等），要求 EN 公开对应的 $f_i(x^*)$ 值。利用拉格朗日插值法，主服务器通过已知的分片 $s(i,j)$ 和公开的 $f_i(x^*)$ 重构多项式片段，并与预提交的哈希承诺 C^i_{hash} 对比，以此验证 EN 是否在整个秘密片段生成过程中保持多项式一致性。若检测到矛盾，主服务器标记该 EN 为异常节点并拒绝其所有秘密片段参与聚合。

（4）安全重构与聚合：SS.recon()

完成分布式验证后，主服务器从至少 t 个 ZS 节点收集同一 EN 的有效片段集合 $\{s_{i,1}, s_{i,2}, \ldots\}$，随后通过拉格朗日插值公式重构标准度 $g_{i,k}$。具体计算中，利用插值点 x_j = ZS 节点编号的特性，可计算分母逆元加速重构过程：

$$g_{i,k} = \sum_{j=1}^{t} \left[s_{i,j} \times \prod_{l \neq j} \frac{(0 - x_l)}{(x_j - x_l)} \right] \tag{13}$$

当所有 EN 的梯度重构完成后，主服务器执行局部聚合操作，计算加权平均梯度：$g_k = \frac{1}{n} \sum_{i=1}^{n} g_{i,k}$。

该协议通过动态门限机制实现弹性安全防护，当 ZSG 覆盖区域内 EN 数量发生变化时，系统可根据固定的安全系统数 λ 动态调整门限值 t。同时，分片元会计算使得 EN 向 mt ZS 发送分片，即使部分 ZS 失效或遭受攻击，仍能通过活跃节点的分片完成梯度重构。

2.3 基于隐私计算的数据安全聚合策略流程

本节基于 2.1 与 2.2 节研究的两种隐私计算方法，提出一种融合轻量化同态加密与动态门限 Shamir 秘密共享的联邦学习数据安全聚合流程。在 FL 每轮训练阶段的初始化阶段，PA-Chain 区块链网络通过分布式密钥生成协议生成全局同态加密公私钥，并依据全局和本区域条件动态

设定秘密共享门限值。边缘节点在收到全局模型和隐私计算配置参数后进行本地训练，并完成模型加密与分片：边缘节点使用 ECC-Paillier 加密本地模型参数，生成密文后通过 TSSS 算法分片并分发给区域服务器。随后，区域服务器验证分片有效性，主服务器收集足够分片后重构区域级加密模型。区块链网络对各区域加密模型执行同态加法聚合，生成全局密文模型。最后，由本轮次的区块链领导节点通过分布式私钥分片协作解密全局密文，获得明文全局模型并更新。该策略通过分阶段加密、分片、验证与聚合，实现跨区域多终端场景下的隐私保护与加密计算。当 FL 轮次为 r 时，聚合策略的流程如算法 1 所示，为区分同态加密和秘密共享的符号变量，用角标 “HE” 表示 ECC-Paillier 同态加密的相关参数，用角标 “SS” 表示动态 TSSS 的相关参数。

Algorithm 1: 基于隐私计算的数据安全聚合策略算法

Input: $|\mathrm{ZSG}| \leftarrow \{\mathrm{ZSG}_1, \mathrm{ZSG}_2, \ldots, \mathrm{ZSG}_l\}$,
EN local model parameters $\leftarrow \{g_{1,1}, g_{1,2}, \ldots, g_{n,l}\}$,
Global model of last round $\leftarrow w_{r-1}$

Output: Updated global model w_r

$(pk, sk_1, \ldots, sk_l) \leftarrow \mathrm{DKG}(t_{\mathrm{HE}}, l)$ // PA-Chain 生成同态密钥
$(\mathrm{GF}(p), S_P, m, t_{\mathrm{SS}}, \mathrm{Hash}()) \leftarrow SS.init()$ // 秘密共享初始化
for k in $(1, \ldots, l)$ **do**
  ZSG_k send $\{k \mid pk, \mathrm{GF}(p), S_P, m, t_{\mathrm{SS}}, \mathrm{Hash}(), w_{r-1}\}$ to EN
  foreach $\mathrm{EN}_{i,k} \in \mathrm{ZSG}_k$ **do**
    $c^i_{\mathrm{HE}} \leftarrow HE.Enc(g_{i,k}, pk)$ // 同态加密生成密文
    $(|s_i|, C^i_{\mathrm{hash}}) \leftarrow SS.share(c^i_{\mathrm{HE}}, T_{ts}, m)$ // 秘密共享生成片段
    Send $\{C^i_{\mathrm{HE}}, |s_i|, C^i_{\mathrm{hash}}\}$ to ZS $\in \mathrm{ZSG}_k$
  foreach $\mathrm{ZS}_{0,k} \in PA\text{-}Chain$ **do**
    $|s_i|_{valid} \leftarrow SS.verify()$ // 秘密共享分布式验证
    $C^k_{\mathrm{HE}} \leftarrow SS.recon(|s_i|_{valid})$ // 秘密共享重构与聚合
$C_{global} \leftarrow HE.add(C^1_{\mathrm{HE}}, \ldots, C^l_{\mathrm{HE}})$ // 同态加法聚合
$sk_{subset} \leftarrow HE.Dec.select(sk_1, \ldots, sk_l, t_{\mathrm{HE}})$
$w_r \leftarrow HE.Dec.threshold(C_{global}, sk_{subset})$ // 同态阈值解密
return w_r

3 安全性理论分析

3.1 正确性分析

定理 1. *若包括边缘节点、区域服务器簇和区块链网络在内的所有 FL 训练参与方按照正确的规则完成数据安全聚合流程，则最终加密聚合结果与明文聚合一致，可得到正确的计算结果，即：*

$$Decrypt\left(\frac{1}{N}\sum_{i=1}^{N} Encrypt(g_i)\right) = \frac{1}{N}\sum_{i=1}^{N} g_i \tag{14}$$

其中 N 为参与训练的节点数，g_i 为第 i 个节点的梯度。

证明. 对于 ECC-Paillier 算法的正确性，设 $\text{Encrypt}(g_1) = (P_1, Q_1)$，$\text{Encrypt}(g_2) = (P_2, Q_2)$，则：

$$P_{\text{sum}} = P_1 + P_2 = \pi_1 G + \pi_2 G = (\pi_1 + \pi_2)G \tag{15}$$

$$Q_{\text{sum}} = Q_1 \cdot Q_2 = (G^{g_1} H^{\pi_1}) \cdot (G^{g_2} H^{\pi_2}) = G^{g_1+g_2} H^{\pi_1+\pi_2} \tag{16}$$

故解密时，有：

$$\text{Decrypt}(Q_{\text{sum}}) = \frac{Q_{\text{sum}}}{P_{\text{sum}}^{sk}} = \frac{G^{g_1+g_2} H^{\pi_1+\pi_2}}{(\pi_1+\pi_2)H} = G^{g_1+g_2} \tag{17}$$

上式得到最终解密结果为 $G^{g_1+g_2}$，即明文 g_1+g_2。因此，$\text{Decrypt}(\text{Encrypt}(g_1)+\text{Encrypt}(g_2)) = g_1+g_2$，即 ECC-Paillier 同态加密算法具有加法同态性。

对于 Shamir 秘密共享方案的正确性，在一个域（通常是有限域）上，给定 t 个不同的点，总存在且仅存在一个次数不超过 $t-1$ 的多项式使得所有点都落在该多项式上，这是多项式插值的基本定理。设多项式 $f(x) = s + a_1 x + \cdots + a_{t-1} x^{t-1}$，分片为 $s_j = f(j)$。由于在构造过程中，分片 s_j 来自于多项式 $f(x)$ 本身，所以用这 t 个点进行插值得到的多项式必定与原多项式 $f(x)$ 完全相同，从而 $f(0)$ 也就必定相等。通过拉格朗日插值，代入 $x = 0$：

$$s = \sum_{j=1}^{t} s_j \cdot \prod_{\substack{1 \le k \le t \\ k \ne j}} \frac{0-k}{j-k} \tag{18}$$

因多项式次数为 $t-1$，插值结果唯一，故重构的 s 与原始密文一致。综合上述两点，加密、分片、聚合与解密过程均为线性操作。设 g_k 为区域主服务器 $\text{ZS}_{0,k}$ 聚合的区域模型参数，则有：

$$\text{Decrypt}\left(\frac{1}{N}\sum_{k}\text{Encrypt}(g_k)\right) = \frac{1}{N}\sum_{k} g_k = \frac{1}{N}\sum_{i=1}^{N} g_i \tag{19}$$

这与联邦学习明文聚合结果一致，由此证明定理 1成立，模型准确性得以保证。 □

3.2 隐私性分析

定理 2. *边缘节点至区域服务器的数据隐私性。假设 ECC-Paillier 加密方案基于 ECDLP 和 Paillier 假设安全，且 Shamir 秘密共享采用 (t, m)-门限机制，则攻击者无法通过截获边缘节点*

传输的密文或分片，恢复本地梯度 g_i，除非其同时满足以下条件：(1) 获取至少 t 个 *Shamir* 分片；(2) 破解 *ECC-Paillier* 加密。

证明. **步骤 1**：设边缘节点使用 ECC-Paillier 加密本地梯度 g_i，生成密文 $C_i = \text{Encrypt}(g_i)$。给定椭圆曲线点 $P = \phi G$，根据 ECDLP 假设，求解随机数 ϕ 的计算复杂度为指数级，即：

$$\Pr[\mathcal{A}(G, P) = \phi] \leq \text{negl}(\lambda) \tag{20}$$

在合数模 n^2 下，基于 DCR 假设，区分 $h^{m_n} n^{\phi} \mod n^2$ 与随机值的概率可忽略，其中 $h = G^{\alpha} \mod n^2$，这里 α 为私钥参数。故存在关系：

$$|\Pr[\mathcal{A}(h^{m_n} n^{\phi} \mod n^2 = 1] - \Pr[\mathcal{A}(U) = 1]| \leq \text{negl}(\lambda) \tag{21}$$

因此，攻击者无法从 C_i 中恢复 g_i，除非同时破解 ECC 的 ECDLP 假设和 Paillier 的 DCR 假设。

步骤 2：将密文 C_i 通过 (t, m)-Shamir 秘密共享分片为 $\{s_{i,1}, \ldots, s_{i,m}\}$。任意 t 个分片可精确重构 C_i，但少于 t 个分片的信息熵为零：

$$H(C_i | \{s_{i,j}\}_{j \in S}) = \begin{cases} 0 & |S| \geq t \\ H(C_i) & |S| < t \end{cases} \tag{22}$$

同时，对任意 $t-1$ 个分片集合 S'，存在一组对应关系，确保 C_i' 多项式满足：$f'(s_{i,j}) = s_{i,j}$ $(j \in S')$ 且 $f'(0) = C_i'$。因此，攻击者无法区分 C_i 与随机密文。

综合上述两个步骤，若攻击者未同时满足获取 t 个分片且破解 ECC-Paillier 加密，则恢复 g_i 的概率可忽略，即：$\Pr[\mathcal{A}(\{s_{i,j}\}, C_i) = g_i] \leq \text{negl}(\lambda)$。这保证了敏感数据在该传输过程中的隐私性，即有效抵御了中间人攻击和推理攻击等主动恶意模型的威胁。 □

定理 3. *抗合谋攻击的能力。若合谋的区域服务器数量为 m'，则当 $m' < t$ 时，无法重构密文；当 $m' \geq t$ 时，需额外破解 ECC-Paillier 加密，其计算复杂度为指数级。*

证明. 若 $m < t$，由定理 1 的分片保护门限特性，少于 t 个分片无法重构密文 C_i，因此合谋服务器无法获得任何有效信息：

若 m' 为合谋服务器，可重构密文 C_i，但仍需解密 C_i 以获得 g_i。根据 ECC-Paillier 的安全性假设，解密需私钥 sk，而私钥通过分布式密钥生成（DKG）分片存储。单个私钥分片 sk_k 的信息熵满足：

$$H(sk \mid sk_k) = H(sk) \tag{23}$$

因此，合谋服务器需同时获取至少 t 个私钥分片（与分片重构独立），其概率为：

$$\Pr[\text{获取 } t \text{ 个私钥分片}] \leq \binom{n}{t}^{-1} \tag{24}$$

当 n 为共识节点总数时，该概率可忽略。故合谋攻击的计算复杂度为：

$$\mathrm{Adv}_{\text{合谋}} \leq \mathrm{negl}(\lambda) + \binom{n}{t}^{-1} \tag{25}$$

这保证了敏感数据在区域服务器处理阶段的隐私性，即有效降低了半诚实服务器发起合谋行为等威胁行为的风险。 □

4 模型性能实验分析

为全面验证 PA-Chain 模型的性能与安全性，实验环境与参数设置如下：

网络模型方面，本实验搭建了 PA-Chain 网络架构来模拟一个标准的跨区域多节点联邦学习场景，具体设置如下：设置 5 个 ZSG，分别模拟不同地理区域的联邦学习场景；每个 ZSG 中设置 5 个区域服务器，每轮处于活跃状态的 ZS 数在 [3, 5] 范围内的整数中随机选择；每个 ZSG 覆盖的区域内随机分布着 10-15 个边缘节点，每轮训练中会有小幅度浮动（±2 个节点），模拟实际场景中的节点动态加入与退出。

数据集和训练模型方面，本实验在图像识别数据集 CIFAR-10 上进行了仿真验证，并选择了 VGG16 模型作为 FL 训练模型，其模型配置与 3.5 节中的定义相同。每个边缘节点在每轮全局训练中进行 5 次局部迭代，全局训练共进行 50 轮次，以确保模型充分收敛。

ECC-Paillier 加密使用了椭圆曲线 secp256k1，密钥长度为 256-bit，密文大小设置为 64 字节。Shamir 秘密共享的安全系数 $\lambda = 0.6$。

为了验证数据安全聚合策略在主动攻击威胁场景下的模型正确性和系统性能开销，本实验模拟了一类攻击情况：

- 主动篡改攻击：每轮内随机选择 10% 的节点作为恶意节点，执行中间人窃听和模型篡改，其中对模型参数进行篡改的威胁模型采用 3.4.1 节定义的 MPA 攻击，篡改检测率通过签名验证与哈希承诺机制进行评估。

4.1 模型准确性分析

为了验证所提出的方法对模型准确性的影响，本节在无任何威胁模型的场景下对所提出的基于隐私计算的数据安全聚合算法（图标中表示为 PA-FL，下同），并与 FedAvg，Paillier 标准同态加密方法和 Shamir 秘密共享算法共三种对照方法进行了比较，模型随训练轮次的准确率变化趋势如图 6 所示。

实验中，将 FedAvg 作为基线方法，无额外的加密、解密的计算和通信开销，因此收敛最快，在约第 35 轮左右达到收敛，最终准确率稳定在 79.2%。而基于 Paillier 的标准同态加密方法因密文膨胀导致计算和通信效率低下，收敛速度显著慢于 FedAvg，约在第 43 轮完成收敛，最终准确率稳定在 74.7%，模型的精确度有一定的损失。而结合动态秘密分片与 ECC-Paillier 同态加密的 PA-FL 聚合策略，收敛略慢于 FedAvg（约第 39 轮），但显著优于 HE（因分片减少通信开销），最终准确率稳定在 77.7%，同经典 Shamir 秘密共享技术得到的最终模型准确率几乎相

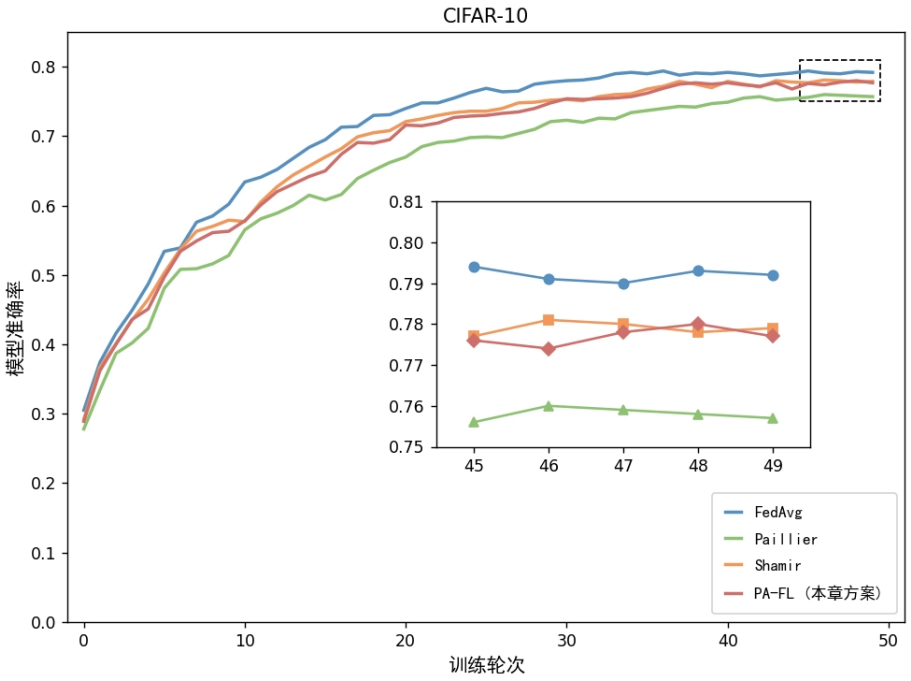

图 6: 在数据集 CIFAR-10 上 10% 恶意节点场景下模型精度随训练轮次变化趋势

表 1: 方案通信开销和计算时长对比

方案	通信开销 (KB/节点)	边缘加密时间 (ms)	区域聚合时间 (ms)
FedAvg	128	0	18
经典 Paillier	512	420	245
静态 Shamir	256	55	105
PA-FL	192	285	135

同，且均接近 FedAvg，用较低的模型准确率损失作为代价得到了隐私性的多层保护，实现隐私与效率的平衡。

4.2 模型性能开销分析

为了验证 PA-FL 策略的通信和计算开销，在数据集 CIFAR-10 上采用文献 [7] 中的数据分布方法，其中边缘节点的训练集规模随机选取 1000-2000 张图像，模型为 VGG-16。对于通信开销，统计单轮训练中边缘节点上传数据量（分片数 × 分片大小 + 签名）。对于计算开销，测量边缘节点加密/分片时间（毫秒）、区域服务器验证与聚合时间（毫秒）。对比方案中，FedAvg 无隐私保护机制，直接传输明文梯度；经典 Paillier 方法使用长度为 2048-bit 密钥，密文膨胀 4 倍；静态 Shamir 采用固定门限 $t_{\mathrm{SS}} = 3$。实验记录的性能开销结果如表 1所示。

由实验结果可得，通信开销方面，经典 Paillier 因密文膨胀（2048-bit 密钥），单节点通信

量高达 512 KB，显著高于其他方案。静态 Shamir 通过分片减少单点传输压力（256 KB），但需冗余分片保证重构。本文方案结合 ECC-Paillier（256-bit 密钥压缩密文）与动态分片（按需调整分片数 m），平均通信量降至 192 KB，较 Paillier 降低 62.5%，较静态 Shamir 减少 25%。

边缘加密时间方面，经典 Paillier 因大整数模幂运算 $O(log^3 n)$，加密耗时 420 ms；本文方案采用 ECC-Paillier（椭圆曲线标量乘法，$O(log\ k)$），加密时间优化至 285 ms，效率较经典 Paillier 提升 32%；静态 Shamir 无加密操作，仅需分片生成（55 ms），但缺乏对分片的隐私保护。

区域聚合时间方面，经典 Paillier 需对密文执行同态加法，聚合耗时 245 ms；静态 Shamir 因固定门限需处理冗余分片（105 ms）；本文方案通过动态门限减少无效分片验证，同时结合轻量化 ECC 运算，使聚合时间降至 135 ms，较 Paillier 减少 45%，较静态 Shamir 优化 28%。

综上，本文提出的联邦学习数据安全聚合策略在跨区域多终端联邦学习场景中，能够以较低的计算与通信代价实现高效的隐私保护，为实际部署提供了理论支撑与技术可行性。

5 结论

为了应对联邦学习在架构中心化和模型参数隐私易泄露上的挑战，本文提出了基于区块链和隐私计算技术的数据安全聚合策略，通过引入隐私计算实现模型数据在传输和计算过程中的全流程隐私保护。首先，研究基于跨区域多终端场景和秘密共享技术搭建了以区域服务器簇为核心设备的隐私聚合链 PA-Chain 网络。其次，在 FL 的本地模型上传和区块链模型聚合过程中引入动态 Shamir 秘密共享和 ECC-Paillier 轻量化同态加密策略，在提高数据安全性的同时增强了 FL 系统在应对隐私推理攻击、半诚实服务器等安全威胁的能力。最终的实验结果表明，本文提出的数据安全聚合策略在保证较小模型精度损失的前提下实现了在通信量、加密和聚合时间上的优化。

参考文献（References）

[1] McMAHAN, Brendan, MOORE, Eider, RAMAGE, Daniel, HAMPSON, Seth, AGUERA Y ARCAS, Blaise. Communication-efficient learning of deep networks from decentralized data [A]. In: Artificial intelligence and statistics [C]. 2017. 1273-1282.

[2] MELIS, Luca, SONG, Congzheng, DE CRISTOFARO, Emiliano, SHMATIKOV, Vitaly. Exploiting unintended feature leakage in collaborative learning [A]. In: 2019 IEEE symposium on security and privacy (SP) [C]. 2019. 691-706.

[3] 赵少飞，吕丽萍，岳剑辉。浅析中间人攻击的手段及防范措施 [J]. 网络安全技术与应用，2023, (12)：21-22.

[4] PAILLIER, Pascal. Public-key cryptosystems based on composite degree residuosity classes [A]. In: International conference on the theory and applications of cryptographic techniques [C]. [出版地未提及，略]. Springer, 1999. 223-238.

[5] GALBRAITH, Steven D. Elliptic curve Paillier schemes[J]. Journal of Cryptology, 2002, 15: 129-138.

[6] SHAMIR, Adi. How to share a secret[J]. Communications of the ACM, 1979, 22(11): 612-613.

[7] CAO, Xiaoyu, FANG, Minghong, LIU, Jia, GONG, Neil Zhenqiang. Fltrust: Byzantine-robust federated learning via trust bootstrapping[J]. arXiv preprint arXiv:2012.13995, 2020.